What is claimed is:

1. A computer implemented method for calculating a normalization factor comprising:

providing a first intensity value ($I^{(1)}$) of a probe in a first probe array and a second intensity value ($I^{(2)}$) of said probe in a second probe array;

obtaining the geometric mean ($x$) of said $I^{(1)}$ and said $I^{(2)}$;

calculating said normalization factor according to:

$f(x) = e^{h(x)}$, wherein said $h(x)$ is derived from referential intensities from said first and second probe arrays.

2. The method of Claim 1 wherein said $h(x)$ is derived by relating geometric means ($x_i'$) of first referential intensities ($RI_i^{(1)}$) in the first probe array and second referential intensities ($RI_i^{(2)}$) in the second probe array to:

$$y_i = \log\left(\frac{RI_i^{(1)}}{RI_i^{(2)}}\right).$$

3. The method of Claim 2 wherein said relating comprising:

sorting ($x_i$, $y_i$) pairs according to $x_i$ into a plurality ($m$ number) of bins with no overlapping;

computing medians ($\bar{x}_k$) of $x_i$*'s* and medians ($\bar{y}_k$) of $y_i$*'s* for each bin; and

interpolating said medians ($\bar{x}_k, \bar{y}_k$).

4. The method of Claim 3 wherein said bins are of approximately equal size.

5. The method of Claim 4 wherein said *h(x)* is:

$$h(x)=\begin{cases}\bar{y}_1, & \text{if } x \le \bar{x}_1\\ w\bar{y}_1+(1-w)\bar{y}_{1+1}, & \text{if } x \in (\bar{x}_1, \bar{x}_1+1), w=\dfrac{\bar{x}_1+1-x}{\bar{x}_1+1-\bar{x}_1}\\ \bar{y}_m, & \text{if } x > \bar{x}_m .\end{cases}\quad ,i=1,\ldots,m-1,$$

6. The method of Claim 5 wherein said $m$ is 3.

7. A computer implemented method for comparing the expression of a gene in a first sample with a second sample comprising:

providing a first plurality of intensity values ($I_i^{(1)}$), each of which reflects the expression of said gene in said first sample, wherein said intensity values are obtained from a first nucleic acid probe array;

providing a second plurality of intensity values ($I_i^{(2)}$), each of which reflects the expression of said gene in said second sample, wherein said intensity values are obtained from a second nucleic acid probe array;

calculating a $p$-value using one-sided Wilcoxon's signed rank test, wherein the $p$-value is for a null hypothesis that *median(f(x)* $I_i^{(2)}$ - $I_i^{(1)}$)=0 and an alternative

hypothesis that $median((f(x)\, I_i^{(1)} - I_i^{(2)})>0$, wherein said $f(x)$ is a normalization factor; and

indicating whether said transcript is present based upon said $p$-value.

8. The method of Claim 7 further comprising a step of calculating normalization factor, said step comprising:

obtaining the geometric mean ($x$) of said $I_i^{(1)}$ and said $I_i^{(2)}$;

calculating said normalization factor according to:

$f(x) = e^{h(x)}$, wherein said $h(x)$ is derived from referential intensities from said first and second probe arrays.

9. The method of Claim 8 wherein said $h(x)$ is derived by relating geometric means ($x_i$') of first referential intensities ($RI_i^{(1)}$) in said first probe array and said second referential intensities ($RI_i^{(2)}$) in said second probe array to:

$$y_i' = \log\left(\frac{RI_i^{(1)}}{RI_i^{(2)}}\right).$$

10. The method of Claim 9 wherein said relating comprising:

sorting ($x_i$, $y_i$) pairs according to $x_i$ into a plurality ($m$ number) of bins with no overlapping;

computing medians ($\bar{x}_k$) of $x_i$'s and medians ($\bar{y}_k$) of $y_i$'s for each bin; and

interpolating said medians ($\bar{x}_k, \bar{y}_k$).

11. The method of Claim 10 wherein said bins are of approximately equal size.

12. The method of Claim 11 wherein said *h(x)* is:

$$h(x) = \begin{cases} \bar{y}_1, & \text{if } x \le \bar{x}_1 \\ w\bar{y}_i + (1-w)\bar{y}_{i+1}, & \text{if } x \in (\bar{x}_i, \bar{x}_i + 1], w = \dfrac{\bar{x}_{i+1} - x}{\bar{x}_{i+1} + -\bar{x}_i} \\ \bar{y}_m, & \text{if } x > \bar{x}_m \,. \end{cases} \quad , i = 1, \ldots, m-1,$$

13. The method of Claim 12 wherein said $m$ is 3.

14. A system for calculating a normalization factor comprising:

a processor; and

a memory coupled with the processor, the memory storing a plurality of machine instructions that cause the processor to perform a plurality of logical steps when implemented by the processor, the logical steps comprising:

providing a first intensity value ($I^{(1)}$) of a probe in a first probe array and a second intensity value ($I^{(2)}$) of said probe in a second probe array;

obtaining the geometric mean ($x$) of said $I^{(1)}$ and said $I^{(2)}$;

calculating said normalization factor according to:

$f(x) = e^{h(x)}$, wherein said *h(x)* is derived from referential intensities from said first and second probe arrays.

15. The system of Claim 14 wherein said $h(x)$ is derived by relating geometric means ($x_i'$) of first referential intensities ($RI_i^{(1)}$) in the first probe array and second referential intensities ($RI_i^{(2)}$) in the second probe array to:

$$y_i = \log\left(\frac{RI_i^{(1)}}{RI_i^{(2)}}\right).$$

16. The system of Claim 15 wherein said relating comprising:

sorting ($x_i$, $y_i$) pairs according to $x_i$ into a plurality ($m$ number) of bins with no overlapping;

computing medians ($\bar{x}_k$) of $x_i$'s and medians ($\bar{y}_k$) of $y_i$'s for each bin; and

interpolating said medians ($\bar{x}_k, \bar{y}_k$).

17. The system of Claim 16 wherein said bins are of approximately equal size.

18. The system of Claim 17 wherein said $h(x)$ is:

$$h(x) = \begin{cases} \bar{y}_1, & \text{if } x \le \bar{x}_1 \\ w\bar{y}_1 + (1-w)\bar{y}_{1+1}, & \text{if } x \in (\bar{x}_1, \bar{x}_1 + 1), w = \dfrac{\bar{x}_1 + 1 - x}{\bar{x}_1 + 1 - \bar{x}_1} \quad, i = 1, \ldots, m-1, \\ \bar{y}_m, & \text{if } x > \bar{x}_m . \end{cases}$$

19. The system of Claim 18 wherein said $m$ is 3.

20. A system for comparing the expression of a gene in a first sample with a second sample comprising:

a processor; and

a memory coupled with the processor, the memory storing a plurality of machine instructions that cause the processor to perform a plurality of logical steps when implemented by the processor, the logical steps comprising:

providing a first plurality of intensity values ($I_i^{(1)}$), each of which reflects the expression of said gene in said first sample, wherein said intensity values are obtained from a first nucleic acid probe array;

providing a second plurality of intensity values ($I_i^{(2)}$), each of which reflects the expression of said gene in said second sample, wherein said intensity values are obtained from a second nucleic acid probe array;

calculating a $p$-value using one-sided Wilcoxon's signed rank test, wherein the $p$-value is for a null hypothesis that $median(f(x)\, I_i^{(2)} - I_i^{(1)})=0$ and an alternative hypothesis that $median((f(x)\, I_i^{(1)} - I_i^{(2)})>0$, wherein said $f(x)$ is a normalization factor; and

indicating whether said transcript is present based upon said $p$-value.

21. The system of Claim 20 further comprising a step of calculating normalization factor, said step comprising:

obtaining the geometric mean ($x$) of said $I_i^{(1)}$ and said $I_i^{(2)}$ ;

calculating said normalization factor according to:

$f(x) = e^{h(x)}$, wherein said *h(x)* is derived from referential intensities from said first and second probe arrays.

22. The system of Claim 21 wherein said *h(x)* is derived by relating geometric means ($x_i$') of first referential intensities ($RI_i^{(1)}$) in said first probe array and said second referential intensities ($RI_i^{(2)}$) in said second probe array to:

$$y_i' = \log\left(\frac{RI_i^{(1)}}{RI_i^{(2)}}\right).$$

23. The system of Claim 22 wherein said relating comprising:

sorting ($x_i$, $y_i$) pairs according to $x_i$ into a plurality (*m* number) of bins with no overlapping;

computing medians ($\overline{x}_k$) of $x_i$*'s* and medians ($\overline{y}_k$) of $y_i$*'s* for each bin; and

interpolating said medians ($\overline{x}_k, \overline{y}_k$).

24. The system of Claim 23 wherein said bins are of approximately equal size.

25. The system of Claim 24 wherein said *h(x)* is:

$$h(x)=\begin{cases}\bar{y}_1, & \text{if } x\le \bar{x}_1\\ w\bar{y}_i+(1-w)\bar{y}_{i+1}, & \text{if } x\in(\bar{x}_i,\bar{x}_i+1],\ w=\dfrac{\bar{x}_i+1-x}{\bar{x}_i+1-\bar{x}_i},\ i=1,\ldots,m-1,\\ \bar{y}_m, & \text{if } x>\bar{x}_m .\end{cases}$$

26. The system of Claim 25 wherein said $m$ is 3.

27. A computer software product for calculating a normalization factor comprising:

computer program code for providing a first intensity value ($I^{(1)}$) of a probe in a first probe array and a second intensity value ($I^{(2)}$) of said probe in a second probe array;

computer program code for obtaining the geometric mean ($x$) of said $I^{(1)}$ and said $I^{(2)}$;

computer program code for calculating said normalization factor according to:

$f(x)=e^{h(x)}$, wherein said *h(x)* is derived from referential intensities from said first and second probe arrays; and

a computer readable medium for storing said codes.

28. The computer software product of Claim 27 wherein said *h(x)* is derived by relating geometric means ($x_i$') of first referential intensities ($RI_i^{(1)}$) in the first probe array and second referential intensities ($RI_i^{(2)}$) in the second probe array to:

$$y_i = \log\left(\frac{RI_i^{(1)}}{RI_i^{(2)}}\right).$$

29. The computer software product of Claim 28 wherein said code for relating comprising:

computer program code for sorting ($x_i$, $y_i$) pairs according to $x_i$ into a plurality ($m$ number) of bins with no overlapping;

computer program code for computing medians ($\bar{x}_k$) of $x_i$'s and medians ($\bar{y}_k$) of $y_i$'s for each bin; and

computer program code for interpolating said medians ($\bar{x}_k$, $\bar{y}_k$).

30. The computer software product of Claim 29 wherein said bins are of approximately equal size.

31. The computer software product of Claim 30 wherein said $h(x)$ is:

$$h(x) = \begin{cases} \bar{y}_1, & \text{if } x \le \bar{x}_1 \\ w\bar{y}_1 + (1-w)\bar{y}_{1+1}, & \text{if } x \in (\bar{x}_1, \bar{x}_1 + 1), w = \dfrac{\bar{x}_1 + 1 - x}{\bar{x}_1 + 1 - \bar{x}_1}, i = 1, \ldots, m-1, \\ \bar{y}_m, & \text{if } x > \bar{x}_m . \end{cases}$$

32. The computer software product of Claim 31 wherein said $m$ is 3.

33. A computer software product for comparing the expression of a gene in a first sample with a second sample comprising:

computer program code for providing a first plurality of intensity values ($I_i^{(1)}$), each of which reflects the expression of said gene in said first sample, wherein said intensity values are obtained from a first nucleic acid probe array;

computer program code for providing a second plurality of intensity values ($I_i^{(2)}$), each of which reflects the expression of said gene in said second sample, wherein said intensity values are obtained from a second nucleic acid probe array;

computer program code for calculating a $p$-value using one-sided Wilcoxon's signed rank test, wherein the $p$-value is for a null hypothesis that $median(f(x)\, I_i^{(2)} - I_i^{(1)})=0$ and an alternative hypothesis that $median((f(x)\, I_i^{(1)} - I_i^{(2)})>0$, wherein said $f(x)$ is a normalization factor;

computer program code for indicating whether said transcript is present based upon said $p$-value; and

a computer readable medium for storing said codes.

34. The computer program code of Claim 33 further comprising computer program code for calculating normalization factor, said code comprising:

code for obtaining the geometric mean ($x$) of said $I_i^{(1)}$ and said $I_i^{(2)}$;

code for calculating said normalization factor according to:

$f(x) = e^{h(x)}$, wherein said *h(x)* is derived from referential intensities from said first and second probe arrays.

35. The computer software product of Claim 34 wherein said *h(x)* is derived by relating geometric means ($x_i'$) of first referential intensities ($RI_i^{(1)}$) in said first probe array and said second referential intensities ($RI_i^{(2)}$) in said second probe array to:

$$y_i' = \log\left(\frac{RI_i^{(1)}}{RI_i^{(2)}}\right).$$

36. The computer software product of Claim 35 wherein said code for relating comprising:

computer code for sorting ($x_i$, $y_i$) pairs according to $x_i$ into a plurality (*m* number) of bins with no overlapping;

computer code for computing medians ($\bar{x}_k$) of $x_i$'s and medians ($\bar{y}_k$) of $y_i$'s for each bin; and

computer code for interpolating said medians ($\bar{x}_k$, $\bar{y}_k$).

37. The computer software product of Claim 36 wherein said bins are of approximately equal size.

38. The computer software product of Claim 37 wherein said *h(x)* is:

$$h(x)=\begin{cases}\bar{y}_1, & \text{if } x \le \bar{x}_1 \\ w\bar{y}_i+(1-w)\bar{y}_{i+1}, & \text{if } x \in (\bar{x}_i, \bar{x}_i+1], w=\dfrac{\bar{x}_{i+1}-x}{\bar{x}_{i+1}+-\bar{x}_i} \quad ,i=1,\ldots,m-1, \\ \bar{y}_m, & \text{if } x > \bar{x}_m .\end{cases}$$

39. The computer software product of Claim 38 wherein said $m$ is 3.